Holstein polaron in the presence of disorder

Mona Berciu,[1] Andrei S. Mishchenko,[2,3] and Naoto Nagaosa[2,4]

[1] *Department of Physics and Astronomy, University of British Columbia, Vancouver, BC, Canada, V6T 1Z1*
[2] *Cross-Correlated Materials Research Group (CMRG), ASI, RIKEN, Wako 351-0198, Japan*
[3] *Russian Research Centre "Kurchatov Institute", 123182 Moscow, Russia*
[4] *Department of Applied Physics, University of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113, Japan*

(Dated: June 6, 2009)

Non-local, inhomogeneous and retarded response observed in experiments is reproduced by introducing the Inhomogeneous Momentum Average (IMA) method to study single polaron problems with disorder in the on-site potential and/or spatial variations of the electron-phonon couplings and/or phonon frequencies. We show that the electron-phonon coupling gives rise to an additional inhomogeneous, strongly retarded potential, which makes instant approximations questionable. The accuracy of IMA is demonstrated by comparison with results from the approximation free Diagrammatic Monte Carlo (DMC) method. Its simplicity allows for easy study of many problems that were previously unaccessible. As an example, we show how inhomogeneities in the electron-phonon coupling lead to nonlocal, retarded response in scanning tunneling microscopy (STM) images.

PACS numbers: 71.38.-k, 72.10.Di, 63.20.kd

Understanding the nature of the materials under the focus of current basic research, as well as the development of novel applications, is unambiguously linked to the physics of quasiparticles in disordered systems and coupled to bosonic modes of different origins. Thus, the manganites which exhibit colossal magnetoresistance are doped materials [1] with considerable electron-phonon (el-ph) as well as electron-magnon and electron-orbitron couplings [2]; the interplay between disorder [3] and the coupling to bosons manifests itself in the peculiarities of their phase diagram [4]. Similarly, the underdoped high temperature cuprate superconductors are inhomogeneous materials [5] with rather strong [6] and inhomogeneous [7] coupling to phonons. Another example concerns charge transport in organic thin-film transistors [8], which is dominated by polaron jumps between different potential traps [9]. The importance of the interplay between disorder and coupling to boson fields is magnified by the fact that weak coupling to a boson mode that is relatively unimportant in a clean system may result in dramatic effects in disordered compounds [10].

The tremendous difficulties in treating the problem of a polaron in the presence of even a single impurity resulted in the 30 year delay between the first results based on the adiabatic approximation [11] and the recent approximation free solution by the DMC method [10]. However, the current level of technology requires theoretical predictions for numerous systems whose inhomogeneity is not limited to a single impurity but includes any form of spatial inhomogeneity, plus potential wells or barriers representing surfaces, interfaces or quantum dot wells, besides cases where both the energy of the bosonic mode and its coupling to the electron are inhomogeneous [12]. Solving such general problems for large systems by numerical methods is still effectively impossible.

In this Letter we study accurately yet efficiently the single Holstein polaron problem with disorder in the potential, the strength of the coupling constant and the frequency of the phonon by developing the Inhomogeneous Momentum Average (IMA) method. The method is based on the Momentum Average (MA) approximation used to study translationally invariant systems, like Holstein [13] and more general models [14]. IMA takes any potential inhomogeneity into account *exactly* and can also handle spatial variations of the coupling constant and of the frequency of the boson modes. Comparing results of IMA with approximation free data from DMC allows us to gauge its accuracy. We perform this comparison in one-dimension (1D) where the worst accuracy of IMA is expected [13]. The IMA approximation can then be systematically improved [15] so that in combination with DMC [10] one gets an accurate and fast tool to study all the systems described above within the framework of a controllable and efficient approximation scheme.

Given the low computational cost of IMA, a rapid scan of large regions of the parameter space is now possible. To compare with experimental observations, we compute STM images of inhomogeneous systems of large enough sizes to render DMC studies impractical, due to the enormous computational cost required for the analytic continuation method [16]. This allows us to prove the nonlocal nature of a system's response to inhomogeneities and demonstrate the importance of the retardation effects.

To avoid cumbersome expressions we consider one impurity in an otherwise homogeneous 1D system, indicating generalizations where suitable. The Hamiltonian is:

$$\mathcal{H} = \mathcal{H}_0 + \hat{U}_0 + \hat{V}_{\rm el-ph} \qquad (1)$$

where $\mathcal{H}_0 = -t\sum_{\langle i,j\rangle}(c_i^\dagger c_j + h.c.) + \Omega\sum_i b_i^\dagger b_i$ is the free part, $\hat{U}_0 = -Uc_0^\dagger c_0$ the on-site attraction to the impurity placed at site $i=0$, and $\hat{V}_{\rm el-ph} = g\sum_i c_i^\dagger c_i\left(b_i^\dagger + b_i\right)$ is the el-ph interaction. The electron's spin is irrelevant.

The goal is to compute the retarded Green's function:

$$G(i,j,\omega) = \langle 0|c_i \hat{G}(\omega) c_j^\dagger|0\rangle = \sum_\alpha \frac{\langle 0|c_i|\alpha\rangle\langle\alpha|c_j^\dagger|0\rangle}{\omega - E_\alpha + i\eta}, \quad (2)$$

where $\hat{G}(\omega) = [\omega - \mathcal{H} + i\eta]^{-1}$ and $\hbar = 1$, because it has information on the single electron eigenstates $\mathcal{H}|\alpha\rangle = E_\alpha|\alpha\rangle$. Also, the local density of states (LDOS) $A(i,\omega) = -\frac{1}{\pi}\mathrm{Im}G(i,i,\omega)$ is measured directly by STM.

We first introduce two additional Green's functions. One is the free electron Green's function

$$G_0(i,j,\omega) = \langle 0|c_i \hat{G}_0(\omega) c_j^\dagger|0\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} dk \frac{e^{ik(R_i - R_j)}}{\omega - \epsilon_k + i\eta}$$

where $\epsilon_k = -2t\cos k$ for nearest neighbor hopping. Generalizations to other dispersions and higher dimensionality are trivial. The second is the "disorder" Green's function $G_\mathrm{d}(i,j,\omega) = \langle 0|c_i \hat{G}_\mathrm{d}(\omega) c_j^\dagger|0\rangle$, corresponding to $\mathcal{H}_\mathrm{d} = \mathcal{H}|_{g=0} = \mathcal{H}_0 + \hat{U}_0$. Using Dyson's identity $\hat{G}_\mathrm{d}(\omega) = \hat{G}_0(\omega) + \hat{G}_\mathrm{d}(\omega)\hat{U}_0\hat{G}_0(\omega)$ straightforwardly leads to:

$$G_\mathrm{d}(i,j,\omega) = G_0(i-j,\omega) - U\frac{G_0(i,\omega)G_0(j,\omega)}{1 + UG_0(0,\omega)}, \quad (3)$$

since $G_0(i,j,\omega) = G_0(i-j,\omega) = G_0(j-i,\omega)$ (the second equality holds if time reversal symmetry is obeyed). Eq. (3) is valid in any dimension, if the appropriate G_0 is used. For more complicated disorder potentials one can find $G_\mathrm{d}(i,j,\omega)$ by a suitable generalization: $\mathcal{H}_\mathrm{d}$ is a quadratic Hamiltonian and can always be diagonalized. Thus, G_d is known and treats the on-site disorder *exactly*.

Since $\mathcal{H} = \mathcal{H}_\mathrm{d} + \hat{V}_{\mathrm{el-ph}}$, we can now proceed to calculate the desired Green's function in terms of $G_\mathrm{d}(i,j,\omega)$. Using Dyson's identity once, we find:

$$G(i,j,\omega) = G_\mathrm{d}(i,j,\omega) + g\sum_{j_1} F_1(i,j_1,\omega)G_\mathrm{d}(j_1,j,\omega) \quad (4)$$

where $F_n(i,j,\omega) = \langle 0|c_i \hat{G}(\omega) c_j^\dagger b_j^{\dagger n}|0\rangle$, with $F_0(i,j,\omega) = G(i,j,\omega)$. Using the Dyson identity again, we find that:

$$\begin{aligned} F_n(i,j,\omega) = {} & g\sum_{j_1\neq j} G_\mathrm{d}(j_1,j,\omega - n\Omega)\langle 0|c_i\hat{G}(\omega)c_{j_1}^\dagger b_{j_1}^\dagger b_j^{\dagger n}|0\rangle \\ & + gG_\mathrm{d}(j,j,\omega - n\Omega)\left[nF_{n-1}(i,j,\omega) + F_{n+1}(i,j,\omega)\right]. \end{aligned} \quad (5)$$

Within the IMA$^{(0)}$ approximation, we set in all these equations $G_\mathrm{d}(j_1,j,\omega - n\Omega) \to 0$ if $j \neq j_1$ and $n > 0$. This is a good low-energy approximation, because the ground-state (GS) of the polaron is below the spectrum of $\mathcal{H}_\mathrm{d}$ and so for $\omega \sim E_{GS}$, $G_\mathrm{d}(i,j,\omega - n\Omega)$ decreases exponentially with the distance $|i-j|$, the decrease being steeper for larger n. In other words, IMA$^{(0)}$ ignores exponentially small terms. Like MA$^{(0)}$, IMA$^{(0)}$ is in fact accurate at all energies because it obeys multiple sum rules [13, 15]. It also becomes exact both for $g \to 0$ and $t \to 0$.

With this approximation only the F_n functions survive in Eq. (5), whose general solution is then $F_n(i,j,\omega) = A_n(j,\omega)F_{n-1}(i,j,\omega)$ [13]. The continued fractions:

$$A_n(j,\omega) = \frac{ngG_\mathrm{d}(j,j,\omega - n\Omega)}{1 - gG_\mathrm{d}(j,j,\omega - n\Omega)A_{n+1}(j,\omega)} \quad (6)$$

are simple to compute. We now insert $F_1(i,j,\omega) = A_1(j,\omega)G(i,j,\omega)$ in Eq. (4), resulting in $G(i,j,\omega) = G_\mathrm{d}(i,j,\omega) + g\sum_{j_1} G(i,j_1,\omega)A_1(j_1,\omega)G_\mathrm{d}(j_1,j,\omega)$. To make this system of coupled equations converge fast with the cutoff in j_1, we define

$$A_n(\omega) = A_n(j,\omega)|_{U=0} = A_n(j,\omega)|_{|j|\to\infty}, \quad (7)$$

since if $U = 0$, then $G_\mathrm{d}(j,j,\omega) = G_0(j,j,\omega) = G_0(0,\omega)$ irrespective of j. The same holds for finite U but $|j| \to \infty$, since sites located very far from the impurity are not sensitive to its presence at the origin. Introducing the "effective interaction" potential:

$$v_0(j,\omega) = gA_1(j,\omega) - \Sigma_{\mathrm{MA}^{(0)}}(\omega) \quad (8)$$

where the bulk MA$^{(0)}$ self-energy is $\Sigma_{\mathrm{MA}^{(0)}}(\omega) = gA_1(\omega)$ [15], we can rewrite the equation for $G(i,j,\omega)$ as:

$$G(i,j,\omega) = G_\mathrm{d}(i,j,\tilde{\omega}) + \sum_{j_1} G(i,j_1,\omega)v_0(j_1,\omega)G_\mathrm{d}(j_1,j,\tilde{\omega}) \quad (9)$$

where $\tilde{\omega} = \omega - \Sigma_{\mathrm{MA}^{(0)}}(\omega)$. This equation is very efficient to solve numerically, because $v_0(j,\omega) \to 0$ rapidly with increasing $|j|$. In fact, a cutoff $|j| \leq 5$ suffices for convergence, although a cutoff of 0 or 1 does *not*, showing that $v_0(j,\omega)$ is spread over a few sites around the impurity. Note that inhomogeneities in the values of g and Ω are easy to deal with, since one simply has to use the appropriate g_j and Ω_j values in Eqs. (6) and (8).

Equation (9) reveals a two-fold role of the el-ph interaction. If the solution was just $G(i,j,\omega) = G_\mathrm{d}(i,j,\tilde{\omega})$, it would mean that the renormalized quasiparticle – the polaron – interacts with the bare impurity potential $\hat{U}_0$. However, the second term shows that the impurity potential is renormalized as well

$$\hat{U}_0 \to \hat{U}_0 + \sum_j v_0(j,\omega)c_j^\dagger c_j, \quad (10)$$

and is no longer local and has significant retardation effects through its ω-dependence. In other words, because of el-ph interactions, the dressed polaron interacts with a renormalized, retarded disorder potential.

Since $\hat{U}_0$ is treated exactly, we expect the validity of this approximation to mirror that of the MA$^{(0)}$ for pure Holstein ($U = 0$) therefore to worsen in lower dimensions, when $g \sim t$ (such that the effective 1D el-ph coupling $\lambda = g^2/(2t\Omega) \sim 1$; remember that the approximation becomes exact for both $\lambda \to 0, \infty$) and for smaller Ω [15]. Even then, a larger U improves the accuracy as it pushes

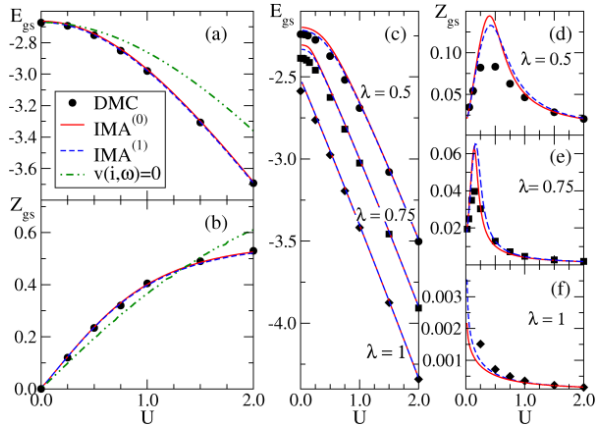

FIG. 1: (color online) (a) and (c) Ground-state energies; and (b),(d)-(f) Spectral weights at the impurity site vs. the impurity potential U, for $t = 1$, $\Omega = 2$ and $g = 1.5$ so that $\lambda = 0.5626$, in (a) and (b), respectively $\Omega = 0.2$ and $g = \sqrt{0.2}, \sqrt{0.3}, \sqrt{0.4}$ so that $\lambda = 0.5, 0.75, 1$ in (c)-(f).

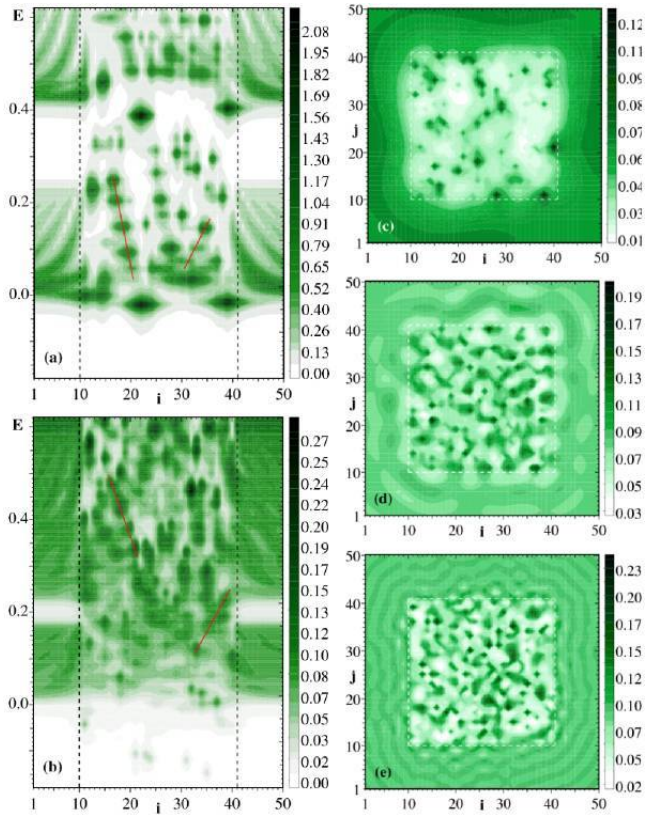

FIG. 2: (color online) (a) 1D LDOS $A(i, E)$. The polaron band and part of the second bound state band are shown. The vertical dashed lines mark the edges of the disordered region. The slanted lines show the dispersion of some features in the spectrum; (b) Analog of (a) for a 2D sample. The plot shows $A(i, j = 25, \omega)$ vs i; (c,d,e) show the 2D $A(i, j, E)$ vs. i, j for $E = 0, 0.25$ and 0.5 respectively. In all plots the energy is measured from the ground-state polaron energy of the corresponding uniform system with $\lambda = 1$, $\Omega = 0.5t$.

the GS to lower energies. However, for small U and Ω and for $\lambda \sim 1$ we need to go to a higher level of the approximation. Like for the homogeneous MA$^{(1)}$ solution [15], in IMA$^{(1)}$ we neglect the exponentially small contributions only if there are $n \geq 2$ phonons present. The IMA$^{(1)}$ solution is similar to Eq. (9), except $\tilde{\omega}$ is now renormalized by the bulk $\Sigma_{\mathrm{MA}^{(1)}}(\omega)$ self-energy [15], while the renormalized potential is also more accurate: $v_1(j,\omega) = g^2 x_{j,\omega} / \left[1 - g x_{j,\omega}[A_2(j,\omega) - A_1(j,\omega-\Omega)]\right] - \Sigma_{\mathrm{MA}^{(1)}}(\omega)$ where $x_{j,\omega} = G_{\mathrm{MA}^{(0)}}(j,j,\omega-\Omega)$ is the MA$^{(0)}$ solution of Eq. (9) at a shifted energy. If necessary, one can go to a higher IMA$^{(n)}$ ($n \geq 2$) level in the same way.

We gauge the accuracy of IMA for model (1) against DMC results. Fig. 1 shows the GS energy and quasiparticle weight Z_{gs} at the impurity site vs. the impurity potential U. The agreement is very good in (a) and (b) even though these are 1D results, where IMA is least accurate. This is partially due to the large $\Omega = 2t$ used [15]. For a worst-case scenario, we plot in (c)-(f) results for a much smaller $\Omega/t = 0.2$, for weak, medium and strong couplings. Now we see clear differences, although they are quantitative, not qualitative. For E_{gs}, the disagreements at small U just mirrors the errors of MA for the Holstein model [15]. As expected, as U increases and E_{gs} moves to lower energies, the accuracy improves. Z_{gs} shows more significant errors at small U, with IMA overestimating the correct answer. This is not surprising, since for such small Ω one expects many phonons to be created at many sites and a higher level n of IMA is needed. However, even levels $n = 0, 1$ capture the physics quite well. Moreover, given the spectral weight sum rules satisfied exactly (6 for IMA$^{(0)}$, 8 for IMA$^{(1)}$, see Ref. [15]), we expect the spectral weight at all energies to be similarly accurate.

In Fig. 1 (a), (b) we also show what happens if we set the additional potential $v_0(j,\omega) \to 0$, i.e. we use an "instantaneous" approximation (see e.g. Refs. [17],[18]) where the el-ph coupling is assumed to renormalize the potential $\hat{U}_0 \to \hat{U}_0 - \frac{g^2}{\Omega}\sum_j c_j^\dagger c_j$. This overall shift is given, in IMA, by the bulk self-energy $\Sigma_{\mathrm{MA}^{(0)}}(\omega \approx E_{\mathrm{gs}}) \approx -g^2/\Omega$ through the renormalized $\tilde{\omega}$. Clearly, this works quite poorly even for this rather small λ, and becomes considerably worse as g and therefore $v_0(j,\omega)$ increase.

In Fig. 2 we show IMA results for a nontrivial problem whose treatment by the DMC method is not feasible due to enormous computational costs. Here we study the site and energy dependent map of the LDOS of an area with randomly chosen el-ph effective coupling $\lambda_i = g_i^2/(2t\Omega) \in [0.9, 1.1]$ embedded in an otherwise uniform infinite system with $\lambda = 1$. We do not add an on-site disorder potential, although it can be included exactly as discussed above. The LDOS maps in Fig. 2(a),(b) show clear evidence of the nonlocal response of the system to such inhomogeneities in the coupling constant, with various features changing their position in space as the en-

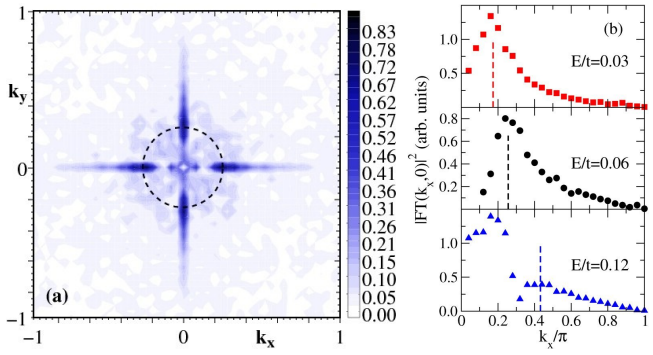

FIG. 3: (color online) (a) Fourier transform of the 2D LDOS map at energy $E = 0.06t$ above the bulk ground-state; (b) Same for $k_y = 0$ and $E/t = 0.03, 0.06, 0.12$. The lines show predicted scattering $\mathbf{k}_s$ vectors. For more details, see text.

ergy is changed and $v_0(j,\omega)$ varies (the slanted lines show some examples). This is more pronounced in 1D, since in 2D some of these shifts proceed in the direction perpendicular to the line with $j = 25$. The retarded nature is apparent not only through the different-looking LDOS maps at different energies inside the disordered region, but also by the slow convergence towards the uniform bulk value and the Friedel-like oscillations seen in the uniform region surrounding it. As expected, the wavelength of these oscillations decreases with increasing energy.

The 2D LDOS maps at different energies show no correlations, but their Fourier transforms (FT) should show peaks at $\mathbf{k}$ values where scattering between *qp* states of that energy is likeliest. Such analysis is well known for STM data in cuprates, for scattering both on impurities [19] and inhomogeneities in the el-ph coupling [7]. In Fig. 3(a) we show the FT at $E = 0.06t$ above the bulk polaron ground-state (we averaged over FT of 200 LDOS maps such as shown in Fig. 2, but for 50x50 site inhomogeneous regions. We removed the $\mathbf{k} = 0$ peak for clarity). At such low energies, the bulk polaron dispersion $E(\mathbf{k}) = -2t^*(\cos k_x + \cos k_y - 2) \approx \hbar^2\mathbf{k}^2/(2m^*)$, so we expect signal up to a $k_s = 2\sqrt{2m^*E/\hbar^2}$ for $\mathbf{k} \to -\mathbf{k}$ scattering. The effective polaron mass is $m^*/m = 1.86$ for $\lambda = 1, \Omega = 0.5t$. We show k_s as a dashed line found to be in agreement with the FT data, as confirmed in Fig. 3(b) for several energies. This demonstrates that it is the polaron and not the bare particle that is scattered by inhomogeneities. We believe that the second peak visible for higher energies is due to inelastic scattering between the first and second polaron bound states, but this needs further study. In any event, these results show agreement with the general phenomenology seen in the experimental data [7, 19] well beyond the single site disorder case usually considered theoretically [20].

Because IMA is very fast (2D LDOS maps such as shown in Fig. 2(c) take about 50s to generate on an ordinary desktop; moreover this type of calculation is ideal to parallelize, with different CPUs for different energies) one can easily study very large disordered regions, for different types of disorder in the on-site potential (whether short- or long-range) and/or el-ph coupling or phonon frequency, in any dimension. Such studies will reveal the effects of each kind of disorder and the way in which multiple types of inhomogeneities do or do not "interfere". Moreover, such studies can be extended to other bare *qp* dispersions as well as models beyond Holstein, such as systems with multiple phonon modes or with electron-phonon coupling dependent on the phonon momentum, where the bulk MA solutions are already known [14].

In conclusion, we have studied the Holstein polaron problem with disorder by developing an accurate, controllable and fast computational method suitable for a huge class of problems that were unaccessible until now.

Acknowledgments: NSERC and CIFAR (M.B.), RFBR 07-02-0067a (A.S.M.), and Grant-in-Aids No. 15104006, No. 16076205, No. 17105002, No. 19048015, and NAREGI Japan (N.N.).

[1] E. Dagotto, T. Hotta and A. Moreo, Phys. Rep. **344**, 1 (2001).
[2] A. J. Millis, Phys. Rev. B **53**, 8434 (1996); V. Perebeinos and P. B. Allen, Phys. Rev. Lett. **85**, 5178 (2000).
[3] M. Uehara et al, Nature **399**, 560 (1999).
[4] Y. Motome, N. Furukawa and N. Nagaosa, Phys. Rev. Lett. **91**, 167204 (2003).
[5] S. H. Pan *et al*, Nature **413**, 282 (2001).
[6] O. Gunnarsson and O. Rösch, J. Phys.: Condens. Matter **20**, 043201 (2008).
[7] J. Lee *et al*, Nature **442**, 546 (2006).
[8] H. Matsui et al, Phys. Rev. Lett. **100**, 126601 (2008).
[9] G. Horowitz and P. Delannoy, J. Appl. Phys. **70**, 469 (1991).
[10] A. S. Mishchenko *et al.*, Phys. Rev. B **79**, 180301 (2009); J. P. Hague, P. E. Kornilovitch and A. S. Alexandrov, Phys. Rev. B **78**, 092302 (2008); N. Pavlenko and T. Kopp, J. Phys.: Condens. Matter **20**, 395203 (2008).
[11] Y Shinozuka and Y. Toyuzawa, J. Phys. Soc. Jpn. **46**, 505 (1979).
[12] A. V. Balatsky and J.-X. Zhu, Phys. Rev. B **74**, 094517 (2006).
[13] M. Berciu, Phys. Rev. Lett. **97**, 036402 (2006); G. L. Goodvin, M. Berciu, and G. A. Sawatzky, Phys. Rev. B **74**, 245104 (2006).
[14] L. Covaci and M. Berciu, Europhys. Lett. **80**, 67001 (2007); *ibid,* Phys. Rev. Lett. **100**, 256405 (2008); G. L. Goodvin and M. Berciu, Phys. Rev. B **78**, 235120 (2008).
[15] M. Berciu, Phys. Rev. Lett. **98**, 209702 (2007); M. Berciu and G. L. Goodvin, Phys. Rev. B **76**, 165109 (2007).
[16] A. S. Mishchenko et al, Phys. Rev. B **62**, 6317 (2000); A. S. Mishchenko, Phys. Usp. **48**, 887 (2005).
[17] E. Burovski, H. Fehske and A. S. Mishchenko, Phys. Rev. Lett. **101**, 116403 (2008).
[18] A. Macridin et al, Phys. Rev. B **69**, 245111 (2004).
[19] K. McElroy *et al*, Nature **422**, 592 (2003).
[20] Q. Wang and D. Lee, Phys. Rev. B **67**, 020511 (2003).